AppliedMicro X-Gene2

Gaurav Singh, Greg Favor, Alfred Yeung

HotChips 2014

Legal Notices

THIS PRESENTATION CONTAINS FORWARD-LOOKING STATEMENTS WITHIN THE MEANING OF U.S. FEDERAL SECURITIES LAWS. THESE FORWARD-LOOKING STATEMENTS, WHICH RELATE TO THE PRODUCTS, OPERATIONS AND BUSINESS PROSPECTS OF APPLIED MICRO, ARE NOT HISTORICAL FACTS, BUT RATHER ARE BASED ON CURRENT ESTIMATES, EXPECTATIONS AND PROJECTIONS. WORDS SUCH AS "WILL," "WOULD," "EXPECTS," "INTENDS," "PLANS," "BELIEVES," "ESTIMATES," AND SIMILAR EXPRESSIONS OFTEN IDENITIFY FORWARD-LOOKING STATEMENTS.

THESE FORWARD-LOOKING STATEMENTS ARE NOT GUARANTEES OF APPLIED MICRO'S FUTURE PERFORMANCE. THEY ARE SUBJECT TO MANY RISKS AND UNCERTAINTIES THAT COULD CAUSE ACTUAL RESULTS TO DIFFER MATERIALLY FROM THOSE CONTEMPLATED BY THE STATEMENTS. PLEASE REFER TO THE COMPANY'S MOST RECENT SEC REPORTS FOR A FURTHER DISCUSSION. NEITHER THE COMPANY NOR THE PRESENTERS ARE OBLIGATED TO UPDATE THESE STATEMENTS OR PUBLICLY RELEASE THE RESULT OF ANY REVISION TO THEM TO REFLECT EVENTS OR CIRCUMSTANCES AFTER THE DATE OF THIS PRESENTATION.

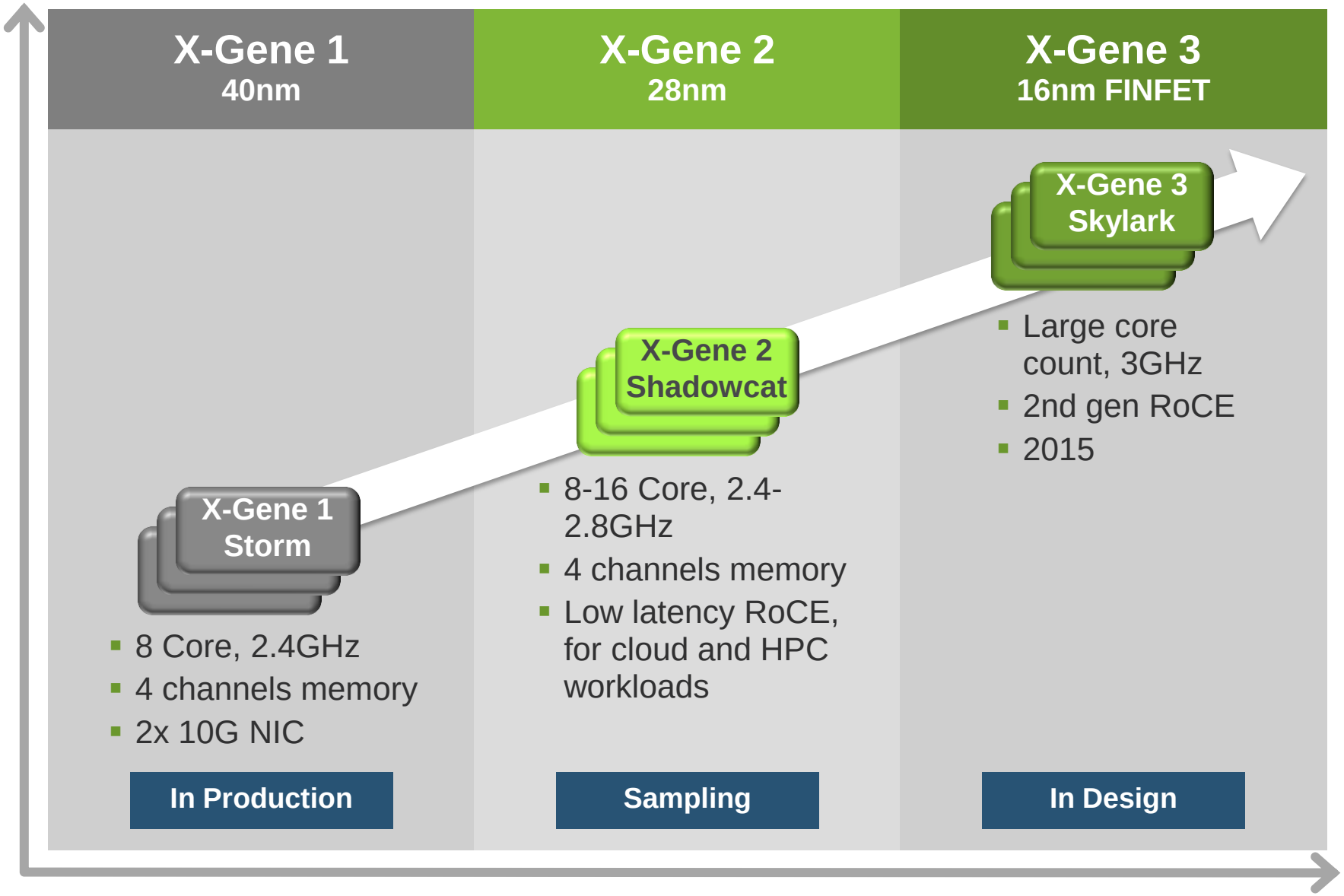
THIS PRESENTATION DOES NOT CONSTITUTE AN OFFER TO SELL OR A SOLICITATION OF AN OFFER TO BUY ANY SHARES OR OTHER SECURITIES OF APPLIED MICRO IN ANY JURISDICTION.

WHILE THE INFORMATION PRESENTED IS BELIEVED TO BE ACCURATE AND COMPLETE AS OF APRIL XX, 2014, NO WARRANTIES OF ANY KIND, EITHER EXPRESS OR IMPLIED, ARE MADE.

Topics

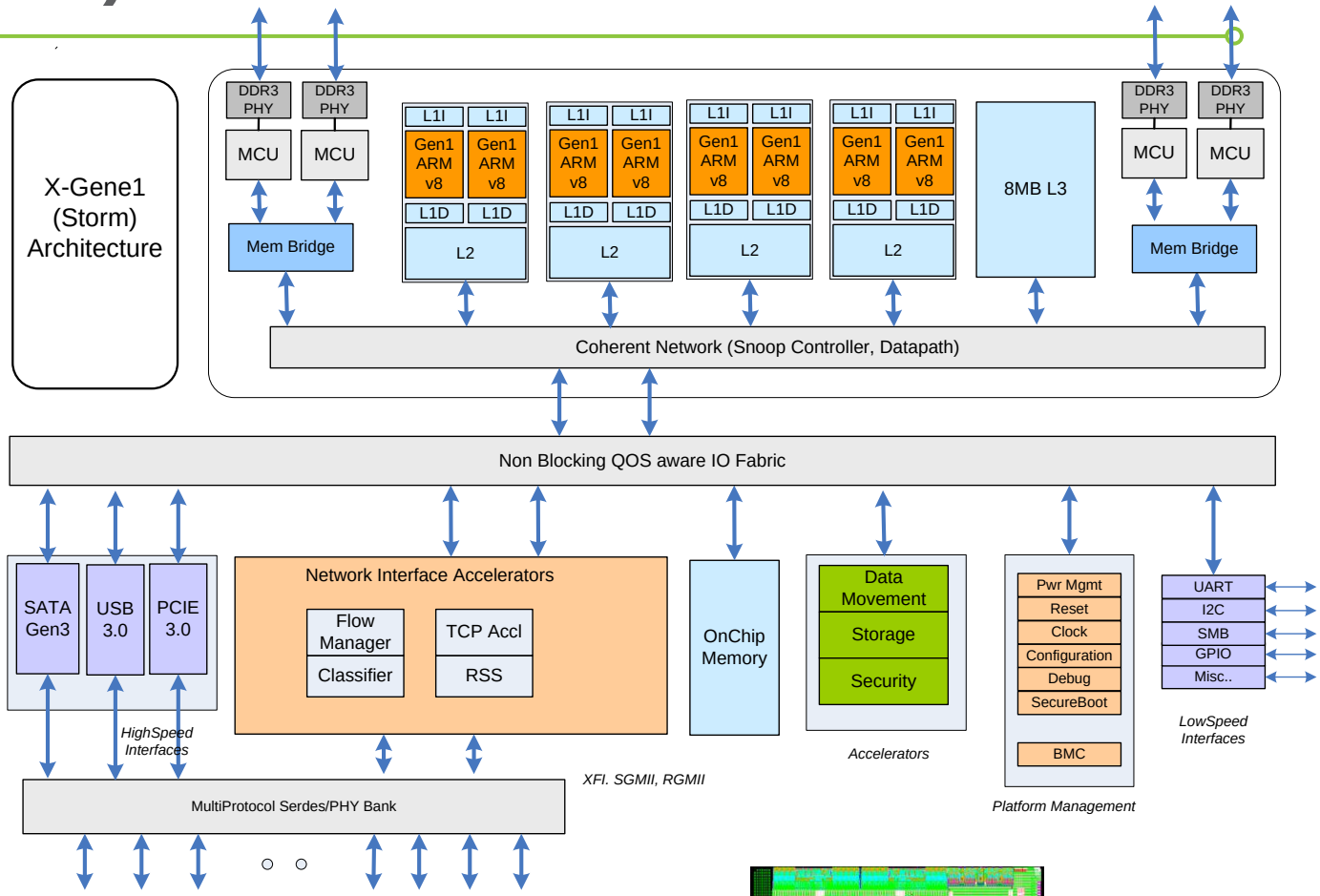
X-Gene	▪ Roadmap
X-Gene1	▪ Review
X-Gene2	▪ Design targets ▪ Architecture ▪ Changes from X-Gene1 ▪ Integrated RoCE HCA ▪ Performance
X-Gene3	▪ Features

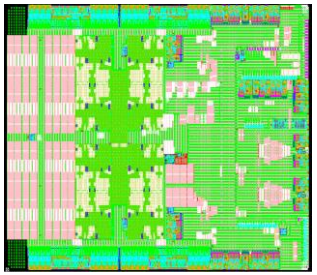
X-Gene: Product Roadmap



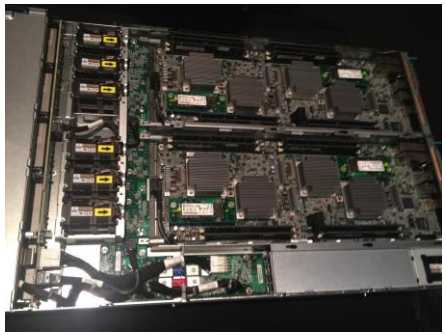
X-Gene1 (Storm)

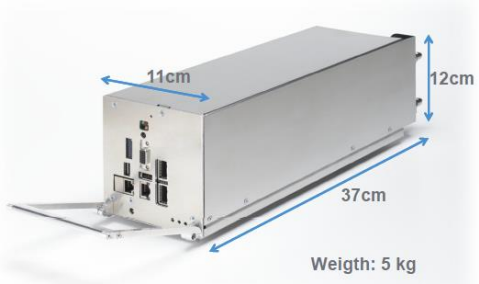
- First generation ARMv8 CPU
- 8 Cores 2.4GHz
- 8MB L3 cache
- 4 DDR memory channels
- PCIE Gen3
- SATA Gen3
- 1/10G interfaces
- Integrated NIC
- 40nm TSMC
- System management
- Security, networking and scale-out accelerators



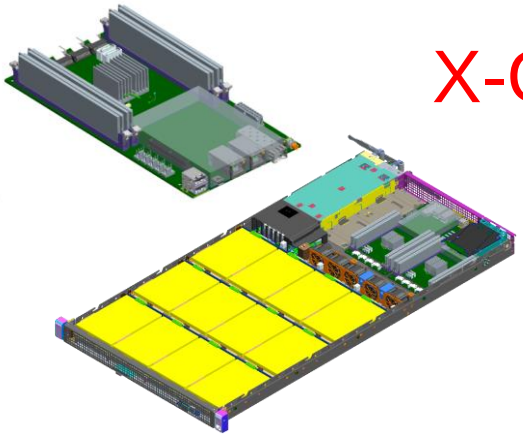


X-Gene 1 Products

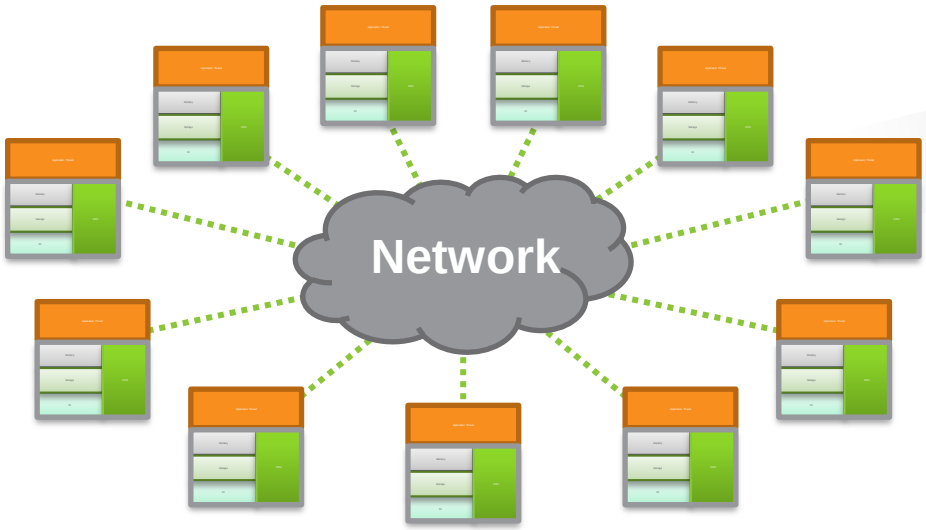


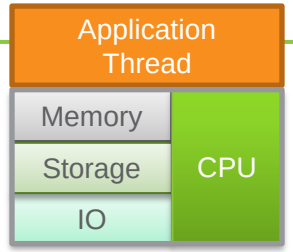


X-Gene ARM Servers are here!



What Are Scale-out Applications





What are scale-out applications?

Pretty much anything that runs in the datacenter today:

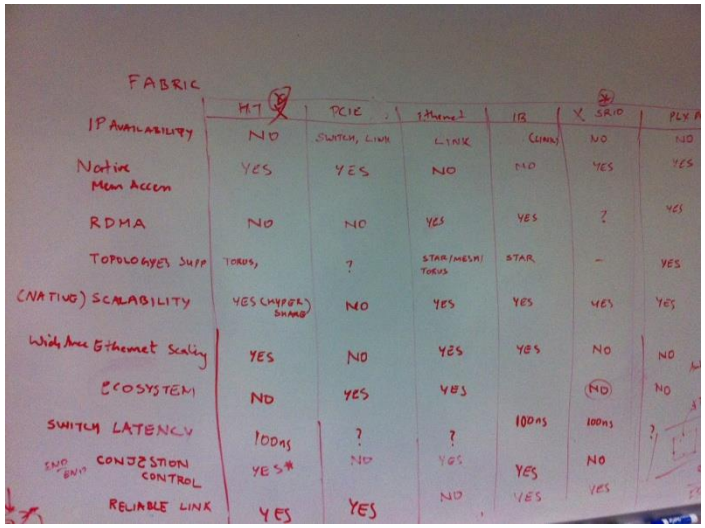
- **Web Front End** ..... Large scale web services hosting
- **Web Search** ..... Index serving, data harvesting
- **Data serving** ..... NoSQL data storage and retrieval
- **Data analytics** ..... Information classification, filtering and extraction
- **Media serving** ..... Large scale hosting and streaming of content

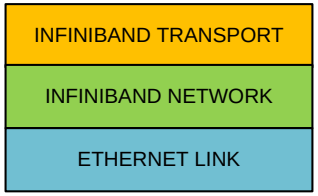
- **They run on 1000s of CPUs. 2P/4P/8P processors do not matter**
- **Communication b/w Application Threads is the bottleneck (Latency, BW)**

- **Thread resources matter**
  - Memory/Thread
  - Memory BW/Thread
  - CacheSize/Thread
  - IO Throughput/Thread
  - Storage Capacity/Thread
- **Application latency**
  - 100s of microseconds
  - TCP/IP as an example:
    - Multiple buffer copies
    - Chatty handshake
    - High s/w overhead

Why RoCE (RDMA over Converged Ethernet)?

- Target the rack level first
  - Let's not change the world .. Yet!
- Use existing switching infrastructure
  - Ethernet is the incumbent
  - Let's not change the link layer
- Reduce latency where it matters
  - Reduce 500ns to 100ns?
  - Or reduce 50us to 4us?
  - Take care of the end-end latency rather than each hop
- Take advantage of existing software infrastructure
  - Standard API interfaces
  - Existing libraries
  - Existing infrastructure





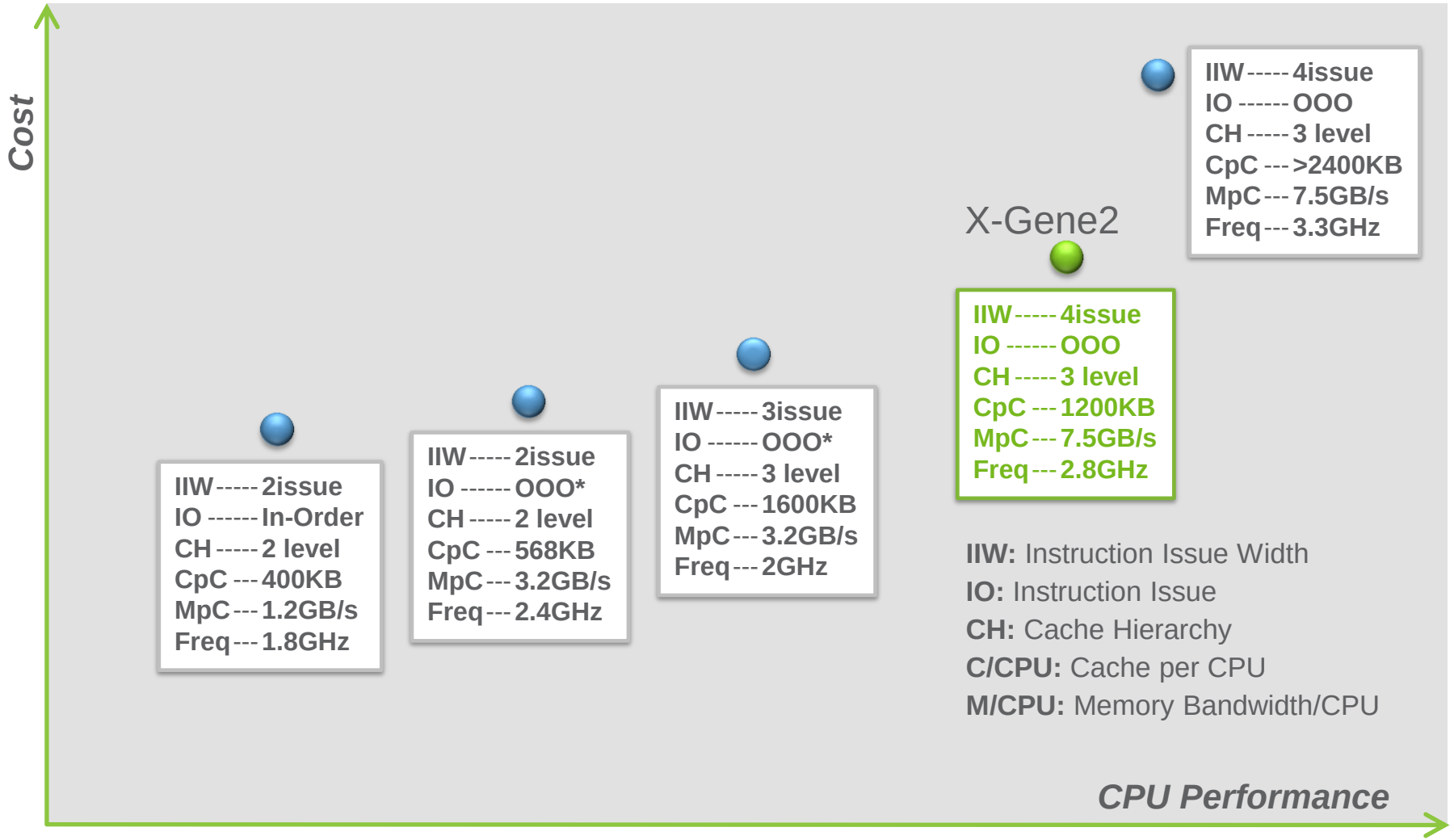
RoCE

	10G Ethernet	10G RoCE
Latency Reduction	1X	6X
Bandwidth Increase	1X	4X

X-Gene2 Target

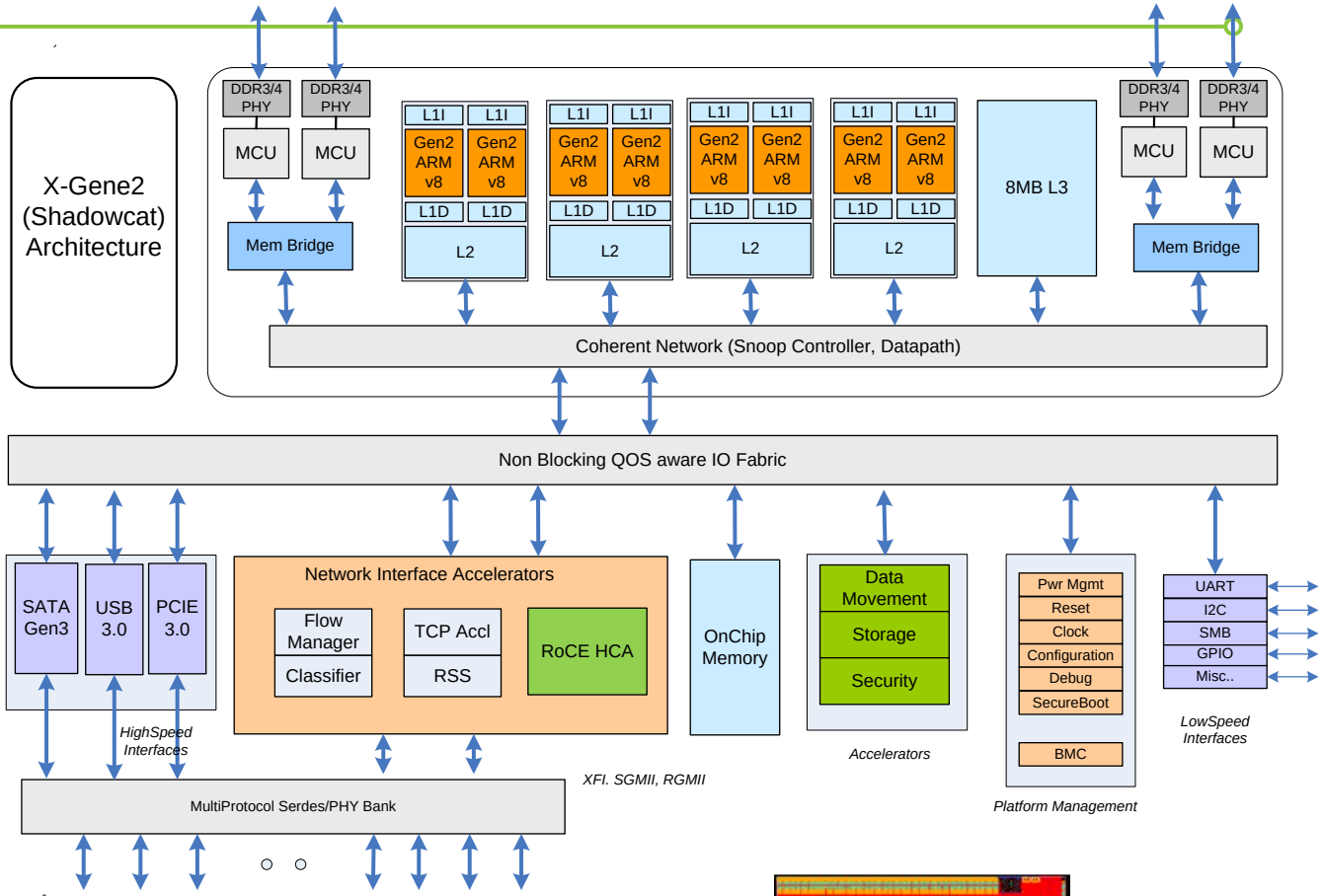
- Design for the rack-level
- Optimize the performance of each application thread
  - **OOO, 4-issue, 2.8GHz CPU**
  - **High memory b/w per thread**
  - **High IO and Storage b/w per thread**
- Design the most efficient application interconnect for these CPUs
- Application interconnect that works not just for 2P/4P, but for hundreds of CPUs
- Reduce application latency by a magnitude of 10
  - **RoCE: 40us → 5us**
- Improve rack level performance at a given latency by 2-3X

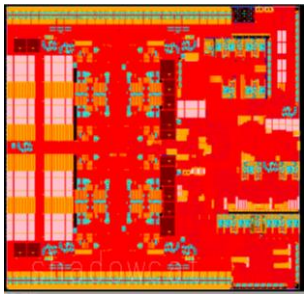
X-Gene2 CPU target



X-Gene2 (Shadowcat)

- Second generation ARMv8 CPU
- 8 Cores 2.8 GHz
- 8MB L3 cache
- 4 DDR channels
- PCIE Gen3
- SATA Gen3
- 1/10G interfaces
- Integrated NIC
- 28nm TSMC
- System management
- Security, networking and cloud accelerators
- Integrated RoCE HCA
- NFV (OVS) and SDN support





X-Gene™ Processor Module

Processor Module

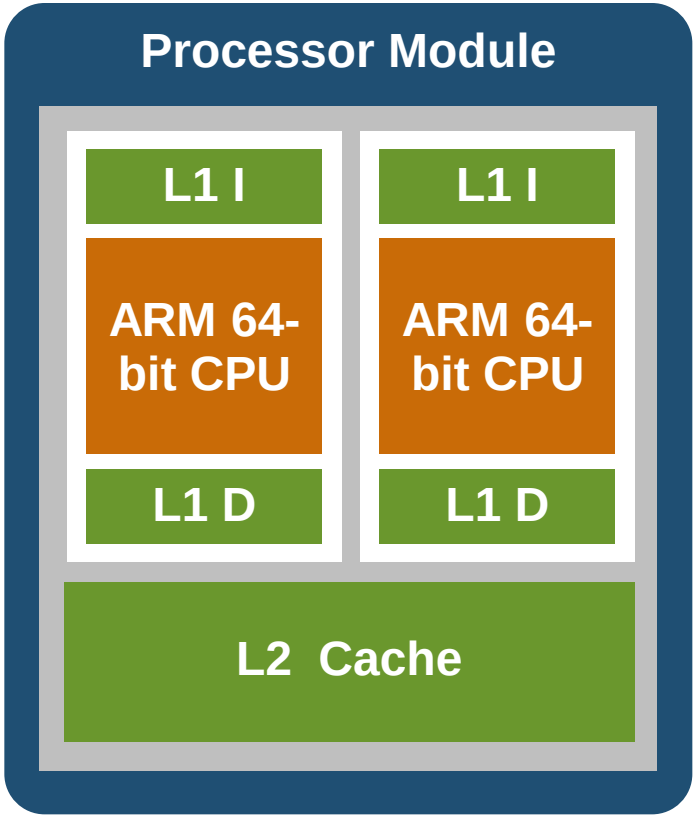
- 2 cores + shared L2 cache
- 4 wide out-of-order superscalar microarchitecture
- >100 instructions in flight
- Integer, scalar, HP/SP/DP FPU and 128b SIMD engine
- Hardware virtualization support
- Hardware tablewalk and nested page tables
- Full set of static and dynamic power management features
  - Fine grain/macro clock gating, DVFS
  - C0, C1, C3, C4, C6 states

Cache Hierarchy

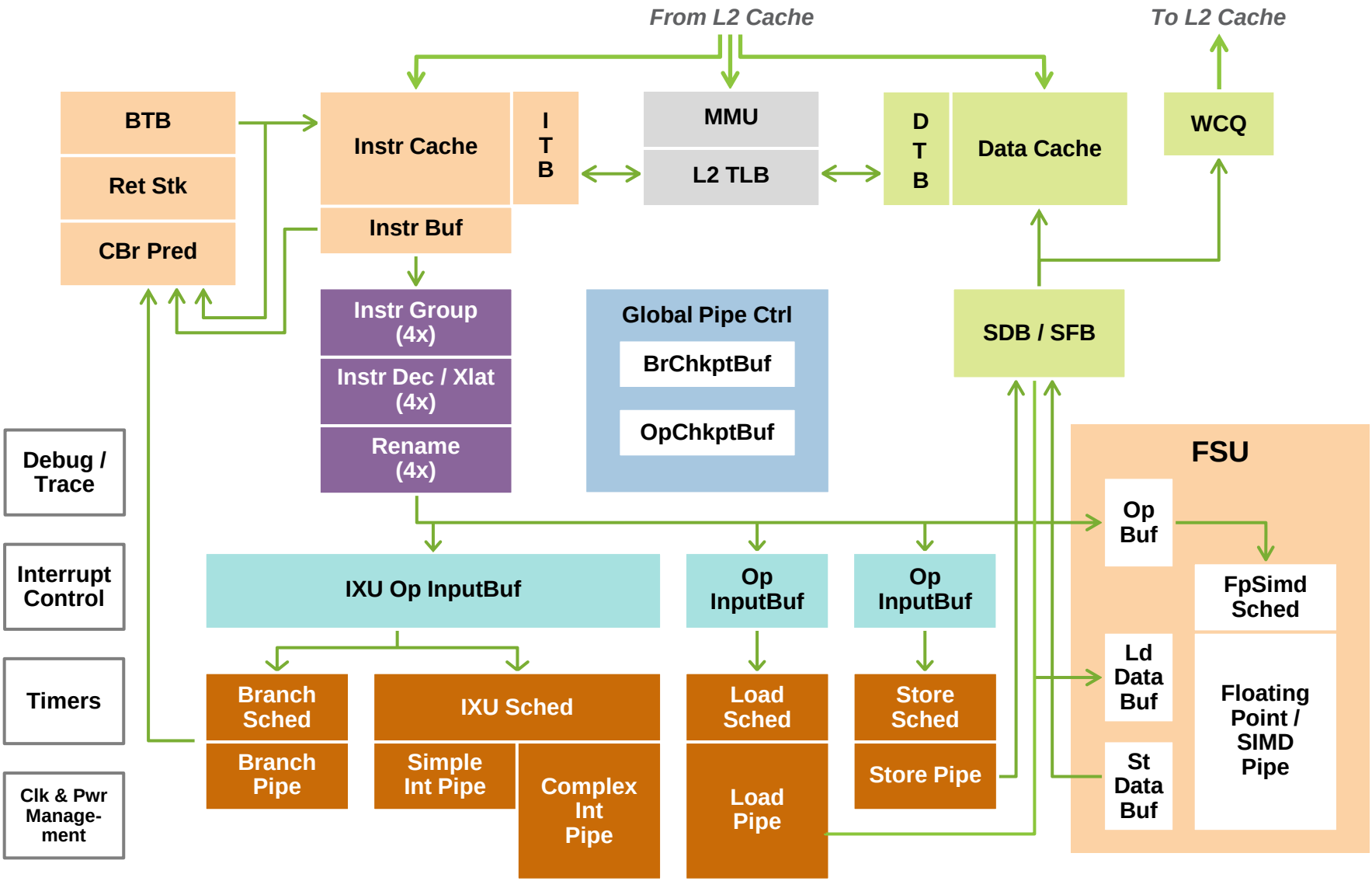
- Separate L1I and L1D caches
- Shared L2 cache among 2 CPUs
- Last-level globally shared L3 Cache
- Advanced hardware prefetch in L1 and L2
- L2 inclusive of L1 write-thru data caches

RAS

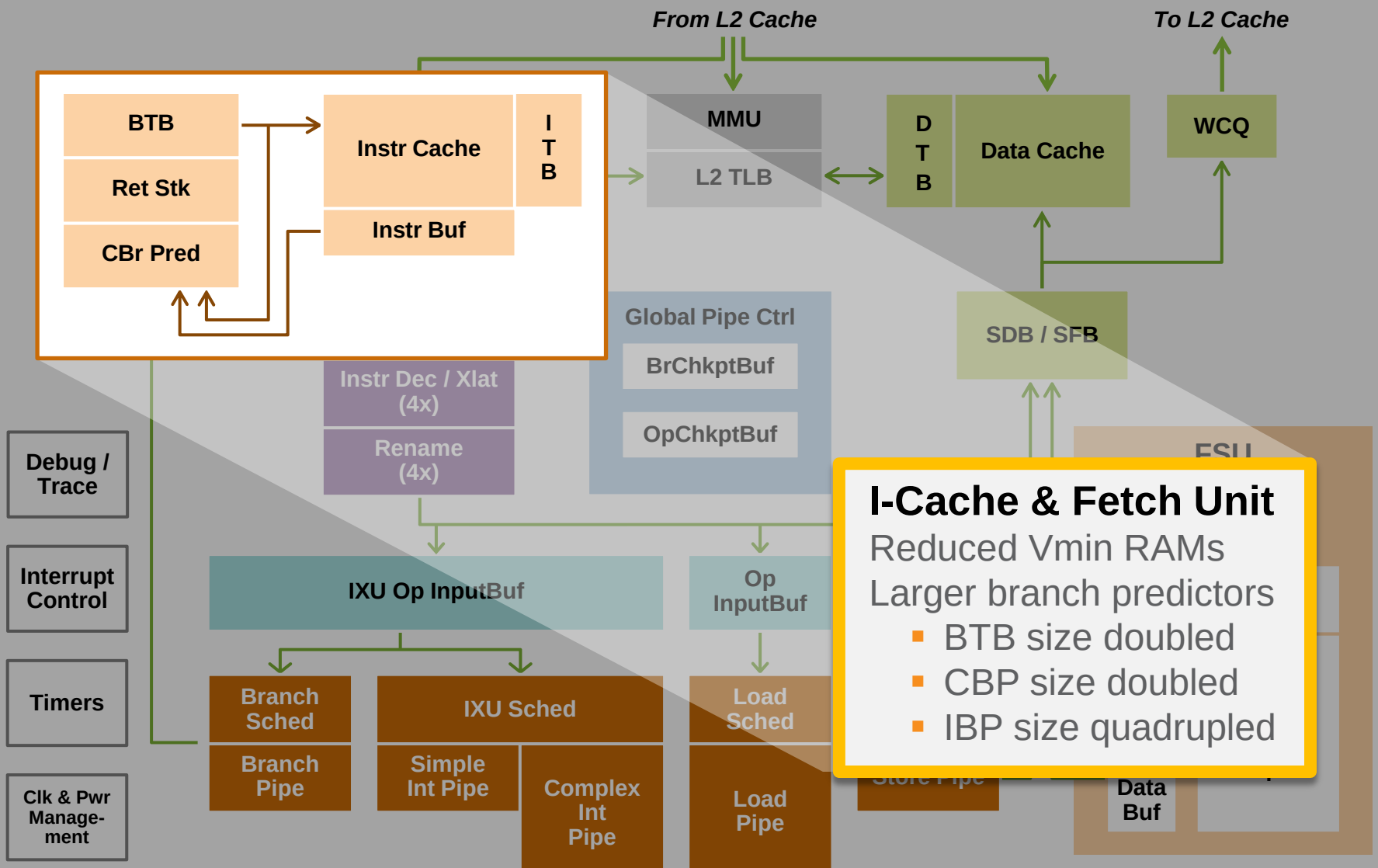
- ECC and Parity protection of all Caches,Tags,TLBs
- Data poisoning and error isolation



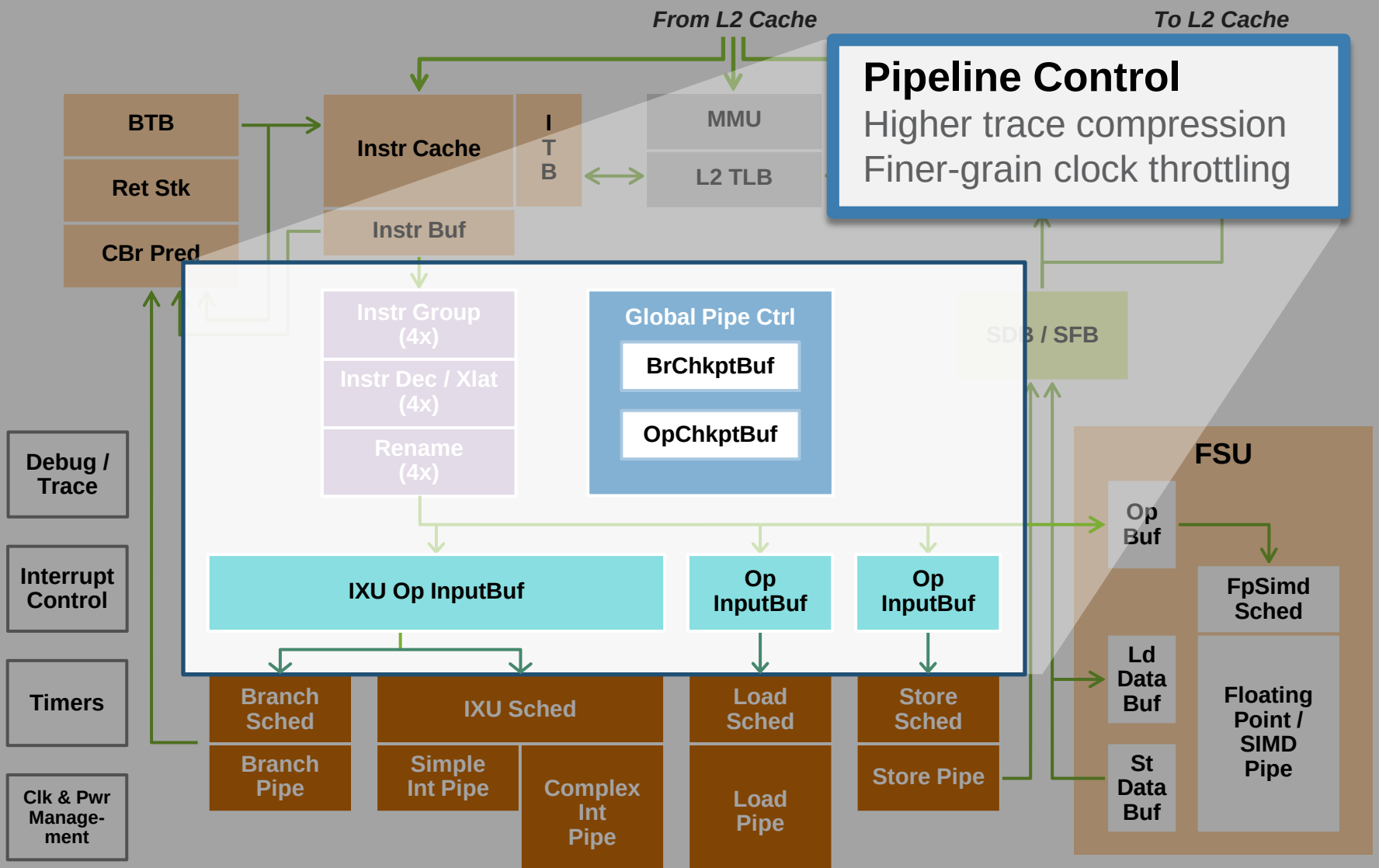
X-Gene2 CPU Block Diagram



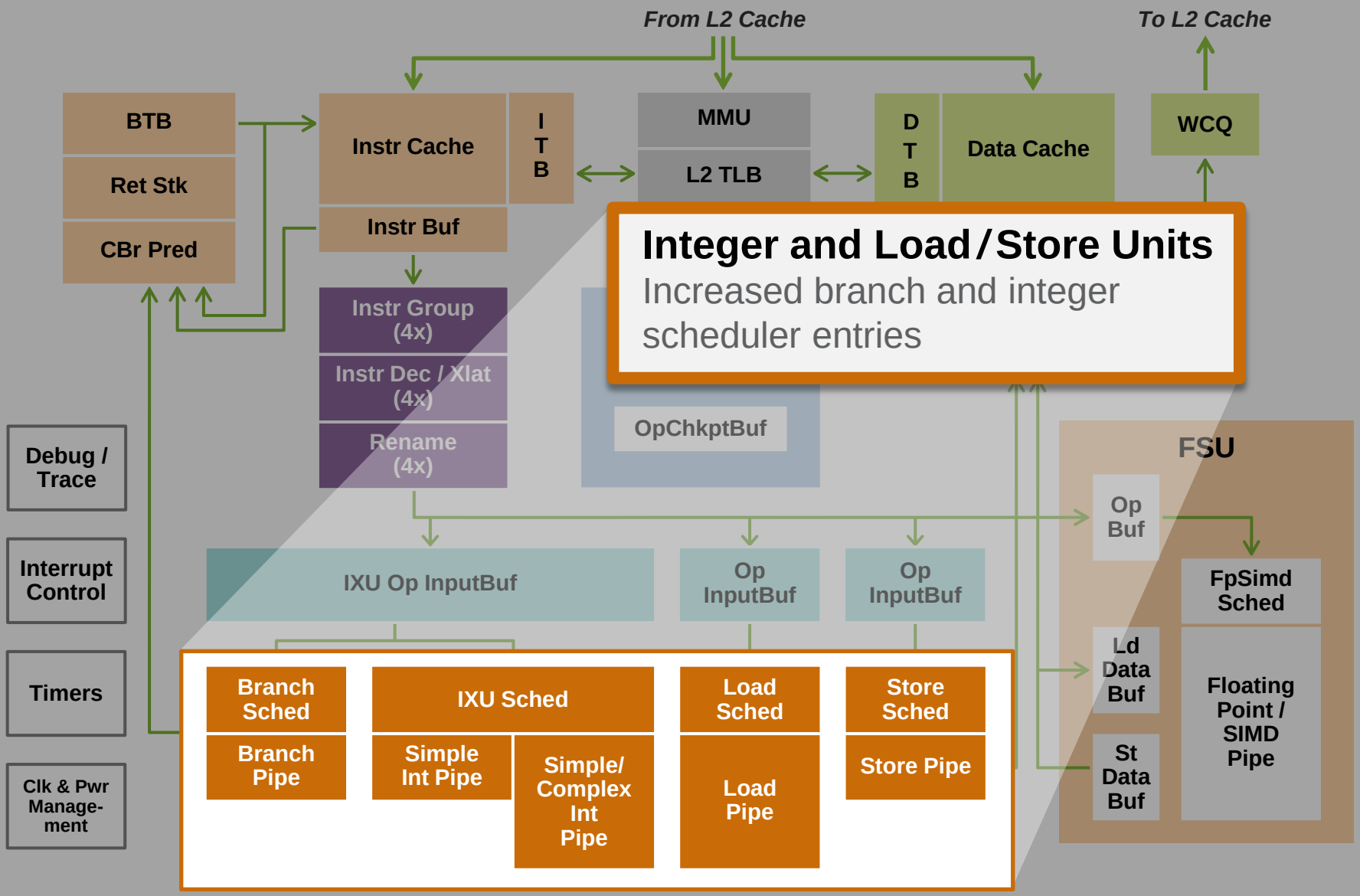
X-Gene2 Instruction Fetch



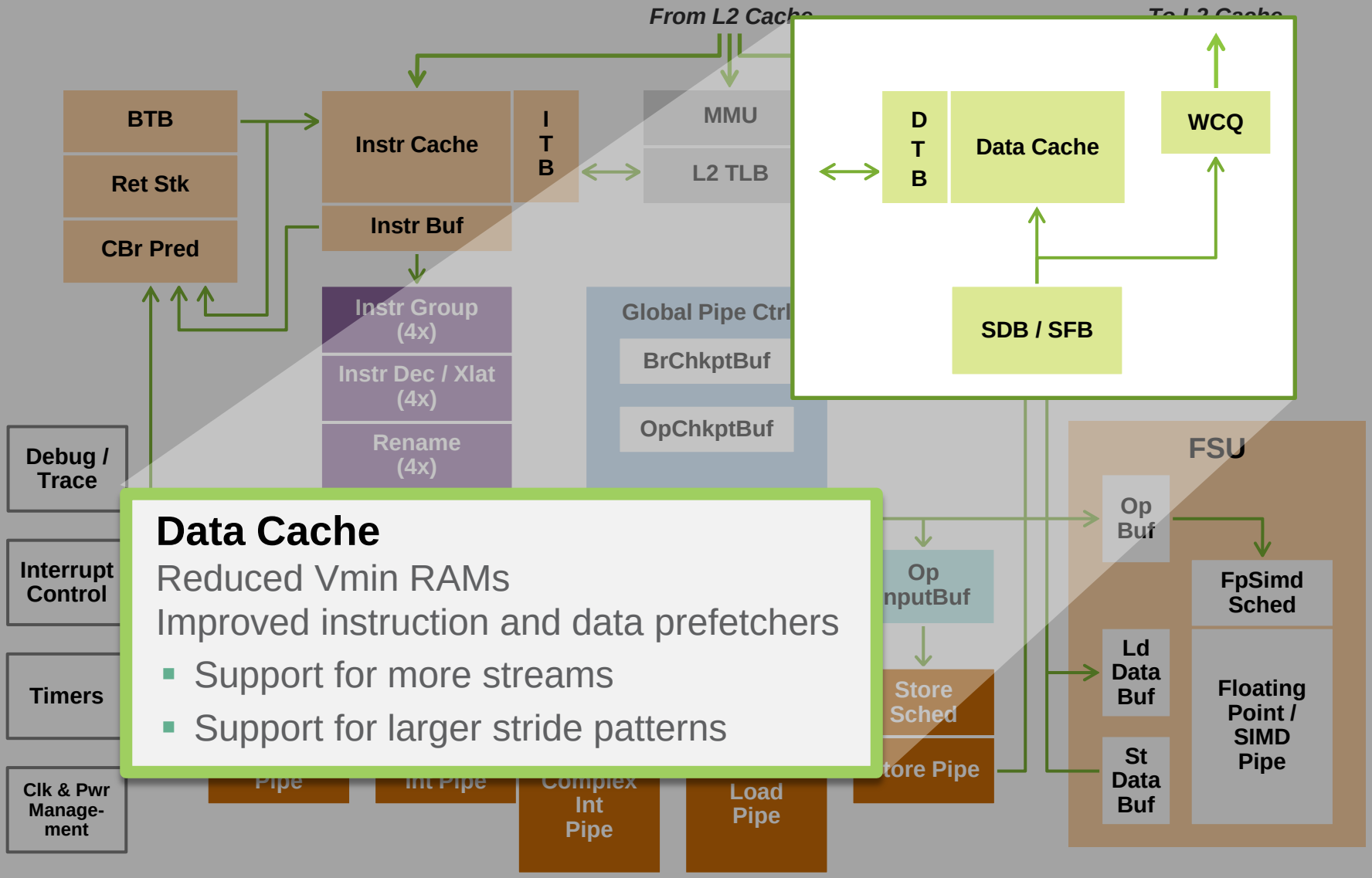
X-Gene2 Reorder Buffer, Dispatch and Control



X-Gene2 Integer, Branch, Load and Store Units



X-Gene2 Data Cache



Data Cache

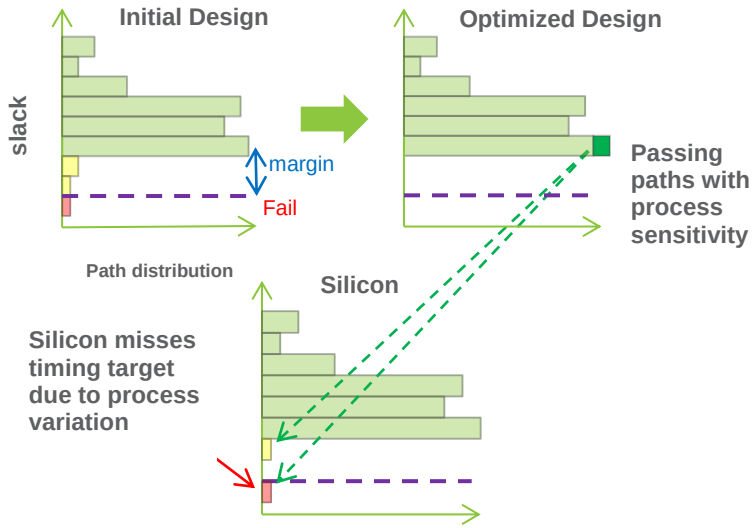
Reduced Vmin RAMs

Improved instruction and data prefetchers

- Support for more streams
- Support for larger stride patterns

Design Challenges from Device Variation

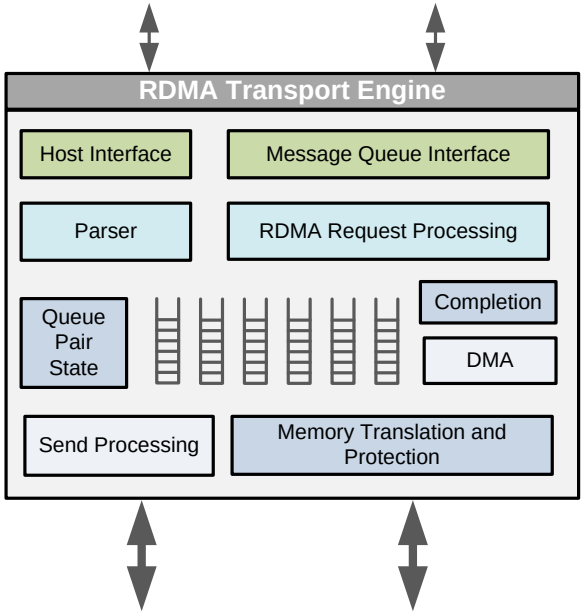
- Impact of device variation is increased in newer technology and increase in transistor count
  - Channel Length may vary due to lithography and optical defects or deviations during etch
  - Vt may vary due to dopant fluctuations or irregularities from the annealing process
  - Oxide thickness variation due to surface roughness scattering
- Additional considerations for process variation required beyond design margining and corner analysis
  - High sigma variation on sensitive gates can dominate a timing path in high frequency designs
  - High number of timing paths lead to an increased likelihood of variation-based outlier paths lowering design frequency
  - Design robustness in memory designs require statistical analysis

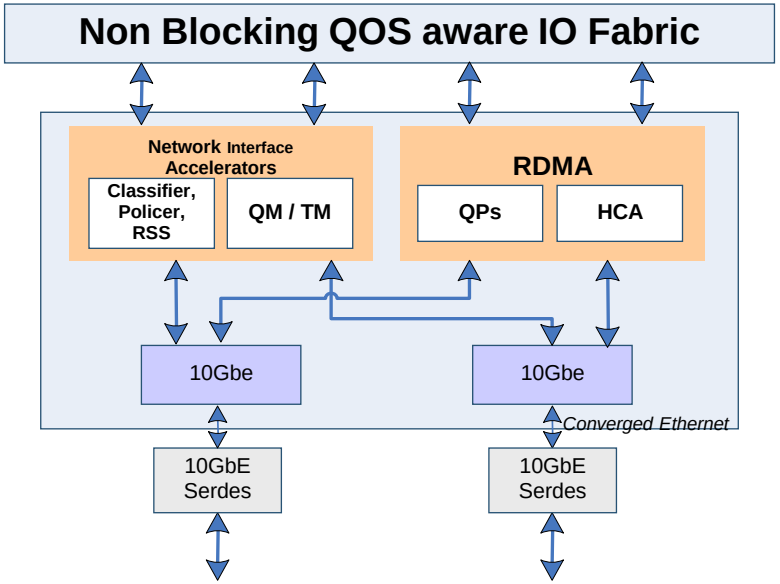


Techniques to address device variation

- P&R level
  - Traditional corner analysis over wide PVT range
  - Design specific OCV margining
  - Cell-based voltage and process sensitivity analysis for top timing paths, which covers thousands of timing paths
  - Usage screening and extensive analysis on hold fixing cells
- Device level
  - Traditional corner analysis over wide PVT range
  - Extensive statistical analysis on all ratio-ed logic, then optimized to obtain near normal distribution
  - Statistical analysis with voltage variance on clock distribution network to limit worst case skew and duty-cycle jitter
  - Device variation optimization on all critical memory timing paths, either with selected devices or carved extraction

RoCE Features in X-Gene2



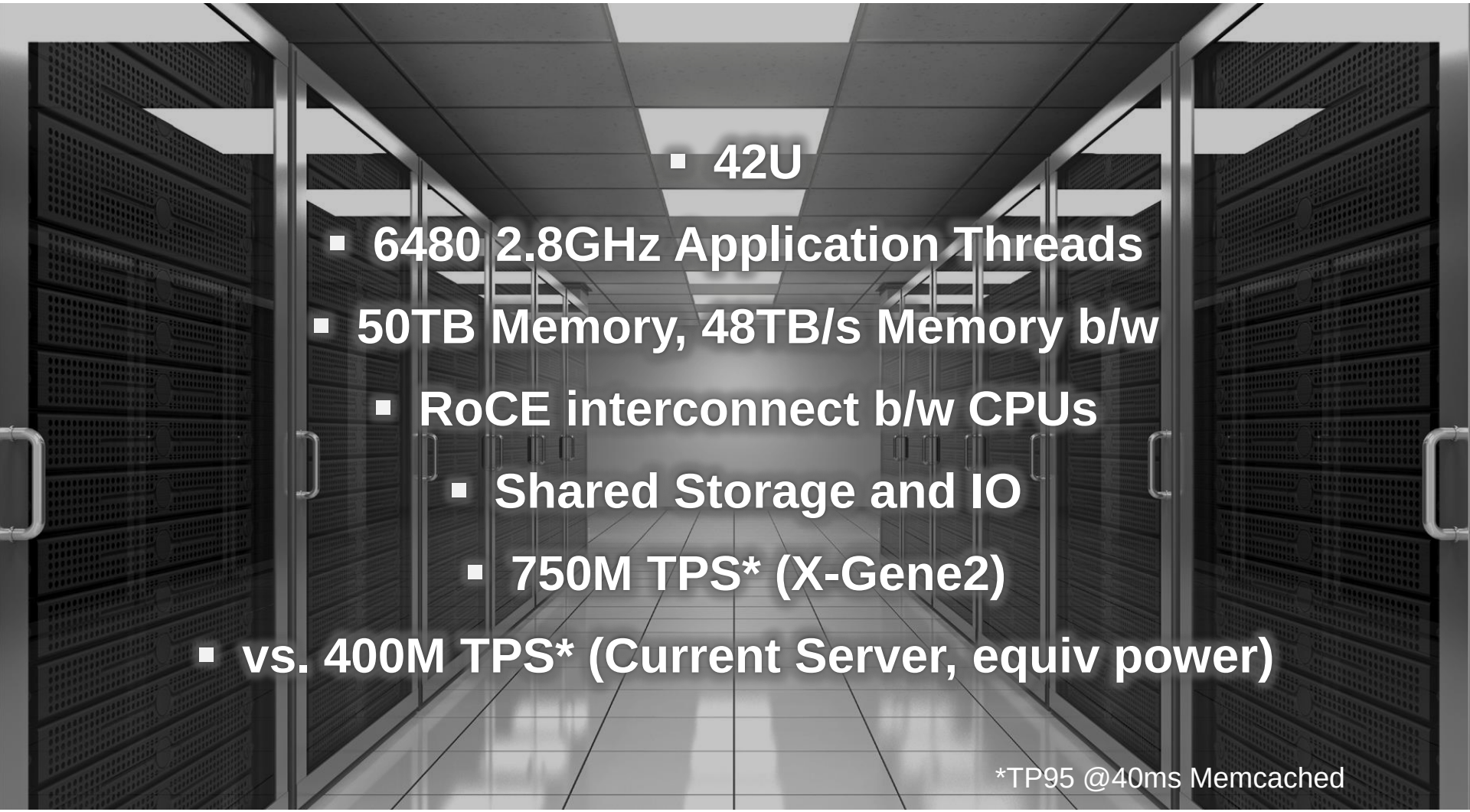


- Verbs Interface compliant
- RDMA over Converged Ethernet
- Hardware-based RC Transport – SRQ support
- RDMA, SEND and ATOMIC operations
- MTU sizes of 256B, 512B, 1KB

- 2 Ports of 10GE
- Large burst of 64B packets and wire-speed processing for < 128B packets
- Along with standard GID, APM specific routable RoCE options

Latency reduction from 50us to 5us compared to TCP/IP

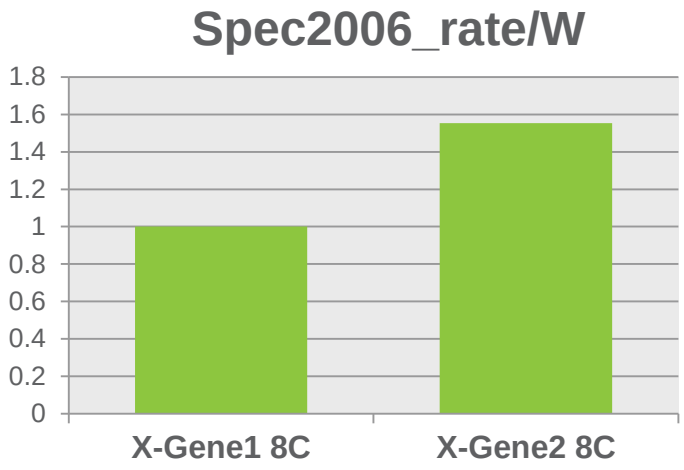
X-Gene2 High Density Rack



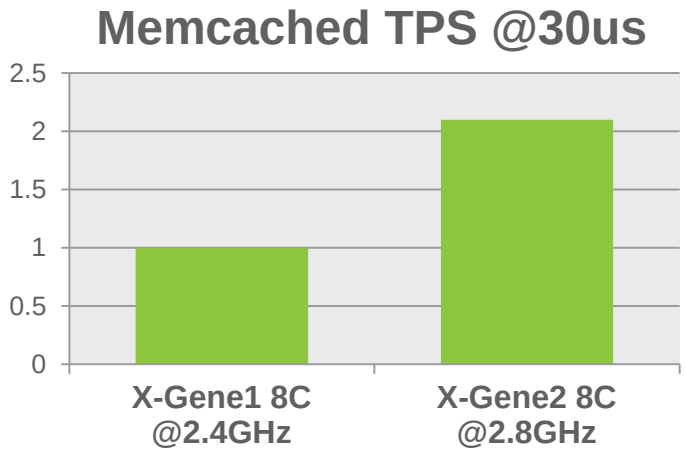
- 42U
- 6480 2.8GHz Application Threads
- 50TB Memory, 48TB/s Memory b/w
- RoCE interconnect b/w CPUs
- Shared Storage and IO
- 750M TPS* (X-Gene2)
- vs. 400M TPS* (Current Server, equiv power)

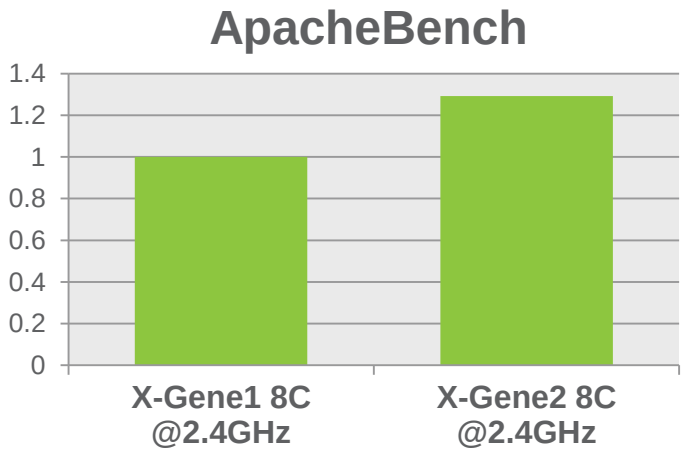
*TP95 @40ms Memcached

X-Gene2 Performance



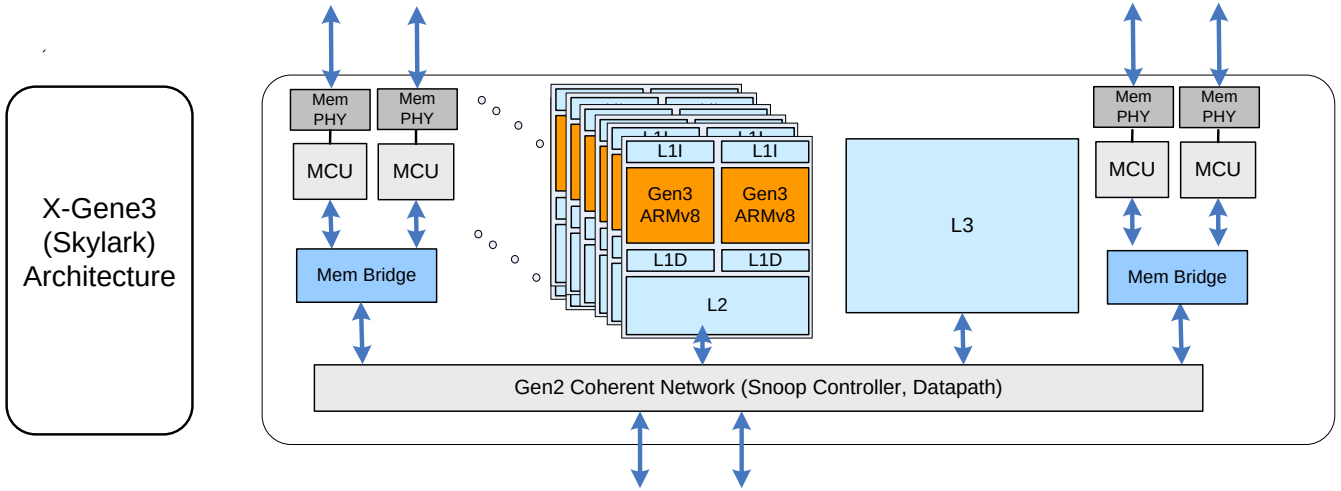
- SpecIntRate2006_Peak w gcc –O3
- Power: TDP





- Performance measured running Apache Web server 2.4.x
- 16KB File request size, 400 concurrent flows

X-Gene3 (Skylark)



- Third generation ARMv8 CPU
- Further micro-architecture enhancements to X-Gene2
- 16-64 Cores 3GHz
- Inter-rack interconnect
- 16nm FinFET
- 2015 samples